Machine Learning Engineer Degree

Capstone Proposal

Madhumathi Prasad
05-05-2018

Domain Background

Food is an important part of everyday life.Now a days people are being conscious about the food they eat.What they eat matters in both nutrition and calorie intake.Not only people who want to improve their health conditions are careful about the food, but many people consider to stay healthy and food eating healthy food has become a part of their lifestyle.

Food recognition is a key element in the food consumption monitoring.There are few traditional techniques in recognizing the food like referring to nutrition facts or Amazon turk.Lately, Deep learning techniques are being used to classify the food. Food recognition is a unique object recognition, it is becoming most active topic in computer vision lately. The specific part is that dish classes have a much higher inter-class similarity and intra-class variation than usual ImageNet objects (cars, animals, rigid objects, etc.).Food are different in different parts of the world;the color,shape and texture makes it interesting to classify them.

Problem statement

How to derive the food information (e.g., food type) from food image effectively and efficiently remains a challenging and open research problem. The goal of the project is to classify food image using CNN.Our model should be able to predict the likelihood of the food image and classify them under the provided categories.Basically it would be a multi-class classification problem in terms of machine learning.

Datasets and Inputs

The dataset which we will be using for this project is 11 categories extracted from Food-101 [5],which has 101 food categories.Each category has around 750 images in training dataset

and 250 images in test dataset.We have further separated training dataset into 7235 images and 1015 images in training and validation datasets respectively.

The food-101 is large dataset and has 1M images,it would take high computational cost and time to evaluate the whole dataset,So We extracted just 11 categories from that dataset.

Solution Statement

As Deep learning techniques have been recognised as one of the best classification problems,in this project we will start with basic CNN technique by adding conv2D, max pooling and dropout layers.Proposed to modify few parameters and optimize to improve accuracy.Would also propose to try transfer learning with pre-trained weights.We have few architecture like RESNET ,vgg-16 and inceptionv3 pretrained on imagenet readily available.Would use few of them to see if it is better in classifying the images under the 11 categories.

Benchmark Model

The benchmark model would be a project from the stanford university,Deep Dish : Deep Learning for Classifying Food Dishes by Abhishek Goswami Microsoft Redmond, WA and Haichen Liu Dropbox Seattle, WA .[6].The model has classified food into 20 categories,which is a different dataset than what we would be using.The model used Transfer Learning (fine tuning a VGG model) with test accuracy of 0.45.This seems to be attainable test accuracy.

Evaluation metrics

We choose test accuracy as our evaluation metric when comparing different models. We would be reporting the test accuracy numbers to two decimal places.The model should be able to predict the images under the provided classes.

The model should be able to predict few images that are not in the dataset under the provided classes.

Project Design

Project Requirements

- Python 3.5
- Tensorflow 1.4.0
- Keras 2.0 with tensorflow as backend.

Convolutional Neural Networks

We will be using Convolutional Neural Networks (CNNs) for this project.What is CNN? CNNs are a special kind of multi-layer neural networks. They are made up of neurons that have learnable weights and biases. Each neuron receives some inputs, performs a dot product and optionally follows it with a non-linearity. The whole network still expresses a single differentiable score function: from the raw image pixels on one end to class scores at the other. And they still have a loss function (e.g. SVM/Softmax) on the last (fully-connected) layer. However, ConvNet architectures make the explicit assumption that the inputs are images, which allows us to encode certain properties into the architecture. These then make the forward function more efficient to implement and vastly reduce the amount of parameters in the network [8].

Project Workflow

- Downloading the dataset and extract 11 categories from 101 categories.Then separate them as train,validation and test dataset.
- Pre processing the dataset by resizing it to make it uniform size.We would also do some data augmentation so we have wide variety of data.
- Building the CNN network with several Convolutional layers followed by pooling layers and dropout layers.Then form fully connected layer with sigmoid activation.
- Evaluating the model with accuracy obtained by the test dataset.The model will be optimised by tweaking hyper-parameters to improve accuracy.
- Extracting features from images of pre-trained model and adding fully connected layers for the 11 categories.

- Evaluating Accuracy and comparing the results.
- Predicting few images which are not in test dataset.

References:

1. Food/Non-food Image Classification and Food Categorization using Pre-Trained GoogLeNet Model Evaluation Metrics.In:https://infoscience.epfl.ch/record/221610/files/madima2016_food_recognition.pdf

2. Food Recognition using Fusion of Classifiers based on CNNs In:https://pdfs.semanticscholar.org/9c1c/4e6ab238f34bef0348b04bfed78cda11e9ed.pdf

3. https://www.analyticsvidhya.com/blog/2017/06/transfer-learning-the-art-of-fine-tuning-a-pre-trained-model.

4. Dataset: https://www.vision.ee.ethz.ch/datasets_extra/food-101/

5. Bossard, L., Guillaumin, M., Van Gool, L.: Food-101 - Mining discriminative components with random forests.

6. http://cs231n.stanford.edu/reports/2017/pdfs/6.pdf

7. Food Image Recognition by Using Convolutional Neural Networks (CNNs).In:https://arxiv.org/pdf/1612.00983.pdf

8. http://cs231n.github.io/convolutional-networks/